TITLE OF THE INVENTION

METHOD AND APPARATUS FOR ESTIMATING A MOTION USING A HIERARCHICAL SEARCH AND AN IMAGE ENCODING SYSTEM ADOPTING THE METHOD AND APPARATUS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001]   This application claims the priority of Korean Patent Application No. 2002-41985, filed on July 18, 2002, in the Korean Intellectual Property Office, the disclosure of which is incorporated herein in its entirety by reference.

BACKGROUND OF THE INVENTION

1.   Field of the Invention

[0002]   The present invention relates to real-time image encoding systems, and more particularly, to a method and apparatus for estimating a motion using a hierarchical search approach, and an image encoding system using the method and apparatus.

2.   Description of the Related Art

[0003]   Among all elements of a general image encoder, a motion estimator performs the most intensive calculations. A fast motion estimation algorithm can reduce a number of calculations performed in the motion estimator. Also, the fast motion estimation algorithm can execute a computation faster than a full search block matching algorithm without degradation in performance. A solution for fast motion estimation is a hierarchical search approach.

[0004]   If the hierarchical search approach is applied to MPEG-2 encoders, a frame motion estimation and a field motion estimation must be performed on all level of images into which a frame image is divided.

[0005]   FIG. 1 is a conceptual diagram illustrating the hierarchical search approach applied to a conventional MPEG-2 encoder. Referring to FIG. 1, a current frame and a reference (or previous) frame form a hierarchical structure through subsampling. In a three-stage hierarchical search approach, the current and reference frames each includes a lowest resolution level (level 2) image 110, an intermediate resolution level (level 1) image 120, and a highest resolution level (level 0) image 130.

[0006]   Hierarchical motion estimation on the hierarchical frame structure begins by initiating a full search on the level 2 image 110 and obtaining initial search points with a minimum Sum of Absolute Differences (SADs). Next, local searches are performed on the level 1 image 120 based on the initial search points obtained from the level 2 image 110. New initial search points with the minimum SADs are extracted by the local searches performed on the level 1 image 120. Thereafter, the local searches are performed on the level 0 image 130 based on the new initial search points obtained from the level 1 image 120. Consequently, the local searches on the level 0 image 130 obtain final motion vectors.

[0007]   Hence, when the MPEG-2 encoder processes prediction frames (P-frames), five hierarchical searches are performed on frames and fields, that is, between frames, on top to top (Top2Top) fields, on top to bottom (Top2Bot) fields, on bottom to top (Bot2Top) fields, and on bottom to bottom (Bot2Bot) fields. When the MPEG-2 encoder processes bidirectional frames (B-frames), a total of 10 hierarchical searches are performed on the B-frames by including forward and backward searches. Accordingly, during the motion estimation in the MPEG-2 encoder, an application of such hierarchical motion searches requires extra memory for use upon both frame and field motion estimations, and demands intensive computation.

SUMMARY OF THE INVENTION

[0008]   The present invention provides a motion estimation method and apparatus to minimize computation during a field motion estimation, while adopting a general hierarchical search approach used in MPEG-2 encoders.

[0009]   The present invention also provides an image encoding system adopting the motion estimation method and apparatus.

[0010]   According to an aspect of the present invention, there is provided a method of estimating a motion of an image of frames organized into a hierarchical structure. In the method, searches are performed on lower level frame data using initial search points to obtain search points with minimum Sum of Absolute Difference (SADs), and the search points with SADs are used as a based motion vector. Then, searches are performed on both upper level frame data and upper level field data using the based motion vector to obtain search points with minimum SADs. The search points with SADs are used as frame and field motion vectors.

[0011]   According to another aspect of the present invention, there is provided an apparatus to estimate a motion of an image of frames organized into a hierarchical structure. The

2

apparatus includes a pre-processor and first and second motion estimation units. The pre-processor performs low-pass filtering and subsampling on a current frame and a reference (previous) frame to organize the current frame and the reference frame into the hierarchical structure. The first motion estimation unit performs a search on frame data obtained by the pre-processor at a low resolution level and searches for at least one initial search point with a minimum Sum of Absolute Difference (SAD). The second motion estimation unit sets initial search points as based frame and field motion vectors and performs a search on the frame data obtained by the pre-processor at a high resolution level by using the based frame and field motion vectors to estimate frame and field motion vectors with minimum SADs.

[0012]   According to an aspect of the present invention, there is provided an apparatus to estimate a motion of an image of frames organized into a hierarchical structure, the apparatus including a discrete cosine transform (DCT) unit performing a discrete cosine transform (DCT) function on the image; a quantization (Q) unit quantizing the DCT function image; a dequantization unit dequantizing the quantized image; an inverse DCT (IDCT) unit performing an inverse discrete cosine transform (IDCT) on the dequantized image; a frame memory (FM) storing the IDCT image on a frame-by-frame basis; a motion estimation (ME) unit forming a hierarchical frame structure by sampling image data of a current frame and the image data of a previous frame stored in the FM, and performing a frame motion estimation by applying based motion vectors (MVs); and a variable length coding (VLC) unit removing statistical redundancy from the quantized image based on the MVs estimated by the ME unit.

[0013]   According to an aspect of the present invention, there is provided a method to estimate a motion of an image of frames organized into a hierarchical structure, the method including: performing a discrete cosine transform (DCT) function on the image; quantizing the DCT function image; dequantizing the quantized image; performing an inverse discrete cosine transform (IDCT) on the dequantized image; storing the IDCT image on a frame-by-frame basis; forming a hierarchical frame structure by sampling image data of a current frame and the image data of a previous frame stored, and performing a frame motion estimation by applying based motion vectors (MVs); and removing statistical redundancy from the quantized image based on the MVs estimated.

[0014]   Additional aspects and/or advantages of the invention will be set forth in part in the description which follows and, in part, will be obvious from the description, or may be learned by practice of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015]   These and/or other aspects and advantages of the invention will become apparent and more readily appreciated from the following description of the aspects of the present invention, taken in conjunction with the accompanying drawings of which:

FIG. 1 is a conceptual diagram illustrating a hierarchical search approach applied to a conventional MPEG-2 encoder;

FIG. 2 is a block diagram of an image encoding system, according to an aspect of the present invention;

FIG. 3 is a block diagram of the motion estimation (ME) unit of FIG. 2;

FIG. 4 is a flowchart illustrating a hierarchical motion estimation method, according to an aspect of the present invention; and

FIG. 5 illustrates an aspect of scaling a based motion vector (BMV) to estimate a motion vector (MV) from top to bottom fields (i.e., a Top2Bot field MV).

DETAILED DESCRIPTION OF THE INVENTION

[0016]   Reference will now be made in detail to the aspects of the present invention, examples of which are illustrated in the accompanying drawings, wherein like reference numerals refer to like elements throughout. The aspects are described below in order to explain the present invention by referring to the figures.

[0017]   Referring to FIG. 2, an incoming image corresponds to a group of picture (GOP). A discrete cosine transform (DCT) unit 220 performs a discrete cosine transform (DCT) function on 8×8 blocks to obtain spatial redundancy from the incoming image.

[0018]   A quantization (Q) unit 230 quantizes the DCT function image. A dequantization unit 250 dequantizes the quantized image.

[0019]   An inverse DCT (IDCT) unit 260 performs IDCT on the dequantized image. A frame memory (FM) 270 stores the IDCT image on a frame-by-frame basis.

[0020]   A motion estimation (ME) unit 280 forms a hierarchical frame structure by sampling image data of a current frame and the image data of a previous frame stored in the FM 270, and performs a motion estimation by applying based motion vectors, obtained from a level 1 image (refer to FIG. 1), to both the frames and fields.

[0021] A variable length coding (VLC) unit 240 removes statistical redundancy from the quantized image based on motion vectors (MV) estimated by the ME unit 280.

[0022] FIG. 3 is a block diagram of the ME unit 280 of FIG. 2 in greater detail. Referring to FIG. 3, a pre-processor 310 performs low-pass filtering (LPF) on the current frame and the reference (previous) frame and organizes each of the current and reference frames in the hierarchical structure through sub-sampling. In a three-stage hierarchical search approach, each of the current and reference frames includes lowest resolution level (level 2) frame data, intermediate resolution level (level 1) frame data, and highest resolution level (level 0) frame data. The highest resolution level (level 0) frame data is the original image, the intermediate resolution level (level 1) frame data has 1/2 of a width and 1/2 of a length of the original image, and the lowest resolution level (level 2) frame data has 1/4 the width and 1/4 the length of the original image.

[0023] A first frame ME unit 320 initiates frame motion estimation by performing a full search on the pre-processed lowest resolution level (level 2) frame data and finding at least one initial search point (motion vector) with a minimum Sum of Absolute Difference (SAD).

[0024] A second frame ME unit 330 continues the frame motion estimation by searching the pre-processed intermediate resolution level (level 1) frame data using the initial motion vectors (MVs), and by finding the initial MVs with the minimum SADs. These initial MVs obtained from the intermediate resolution level (level 1) frame data are referred to as based MVs.

[0025] A third frame ME unit 340 involves both the frame motion estimation and the field motion estimation by performing local searches on the highest resolution level (level 0) frame data, using the based MVs. Then, the third frame ME unit 340 finds frame MVs with the minimum SADs, and field MVs between identical fields that have the minimum SADs. The MVs from a current top field to a previous top field are referred to as Top2Top field MVs. The MVs from the current top field to a previous bottom field are referred to as Top2Bot field MVs. The MVs from a current bottom field to the previous top field are referred to as Bot2Top field MVs. The MVs from the current bottom field to the previous bottom field are referred to as Bot2Bot field MVs. Extra computations are not required to obtain the Top2Top and the Bot2Bot field MVs because the Top2Top and the Bot2Bot field MVs can be obtained based on inter-field SADs automatically obtained upon frame motion estimation.

[0026] A first scaler 350 scales the based MVs using a distance between the top field of the current frame and the bottom field of the previous frame.

[0027] A first field ME unit 360 performs field motion estimation with local searches between the top field of the current frame and the bottom field of the previous frame. The searches occur at the highest resolution level (level 0) using the based MVs scaled by the first scaler 350 and the Top2Bot field MV is obtained with the minimum SAD.

[0028] A second scaler 370 scales the based MVs using the distance between the bottom field of the current frame and the top field of the previous frame.

[0029] A second field ME unit 380 performs the field motion estimation with local searches on the bottom and top fields of the current and previous frames, respectively. The searches occur at the highest resolution level (level 0) using the based MVs scaled by the second scaler 370 and the Bot2Top field MV is obtained with the minimum SAD.

[0030] FIG. 4 is a flowchart illustrating a hierarchical motion estimation method, according to an aspect of the present invention. First, the current frame and the previous frame are each organized in the hierarchical structure by the LPF and the subsampling.

[0031] Next, at operation 410, the frame motion estimation occurs with a full search at the lowest resolution level (level 2), extracting at least one initial search point (i.e., initial MV) with the minimum SAD.

[0032] Thereafter, at operation 420, the frame motion estimation occurs with local searches using the extracted initial MVs at the intermediate resolution level (level 1), obtaining a frame MV with a minimum SAD. A frame-by-frame MV obtained at the intermediate resolution level (level 1) is referred to as the based MV.

[0033] Then, at operations 430 and 440, the frame motion estimation and the field motion estimation occur with local searches at the highest resolution level (level 0) using the based MV, estimating both the frame MVs with the minimum SADs and the field MVs with the minimum SADs.

[0034] The Top2Top and Bot2Bot field MVs are estimated with reference to inter-field SADs automatically obtained upon frame motion estimation. During the Top2Bot and Bot2Top field motion estimations, the based MV cannot be applied without change, because the distance between top and bottom fields differs from the distance between the top fields and between the bottom fields. Accordingly, the Top2Bot and Bot2Top field motions are estimated by local searches based on a new based MV, scaled in consideration of the distance between

corresponding fields. Through the local searches, the Top2Bot and Bot2Top field MVs with the minimum SADs are obtained.

[0035]   To sum up, instead of performing searches at the lowest and intermediate resolution levels (levels 2 and 1) to achieve a field motion estimation, local searches are performed at the highest resolution level (level 0) using the search points (i.e., MVs) obtained by the frame motion estimation at the intermediate resolution level (level 1).

[0036]   FIG. 5 illustrates an aspect of scaling the based MV to estimate the MV between the bottom and top fields (i.e., a Top2Bot field MV). Referring to FIG. 5, it is assumed that each current and previous frame has the top field and the bottom field. If a distance between identical fields (e.g., between top fields or between bottom fields) is m, a distance between different fields (e.g., between bottom and top fields) is n, and a based MV is BMV, the Top2Bot field MV can be estimated as BMV×n/m. In addition, an offset, which considers field characteristics, is added to the scaled BMV. A Bot2Top field MV can be obtained in a way similar to the estimation of the Top2Bot field MV.

[0037]   While the present invention has been particularly shown and described with reference to exemplary aspects thereof, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention as defined by the following claims. The invention can also be embodied as computer readable codes on a computer readable recording medium. The computer readable recording medium is any data storage device that can store data, which can be thereafter read by a computer system. Examples of the computer readable recording medium include read-only memory (ROM), random-access memory (RAM), CD-ROMs, magnetic tapes, floppy disks, optical data storage devices, and so on. Also, the computer readable code can be transmitted via a carrier wave such as the Internet. The computer readable recording medium can also be distributed over a computer system network so that the computer readable code is stored and executed in a distributed fashion.

[0038]   As described above, when an MPEG-2 encoder adopts a hierarchical search approach, a based MV obtained from an intermediate resolution level (level 1) is applied to a frame motion estimation and a field motion estimation, thus, reducing an amount of computation required for motion estimation.